PERGAMON

Vision Research 43 (2003) 205–219

Vision Research

www.elsevier.com/locate/visres

The selective tuning model of attention: psychophysical evidence for a suppressive annulus around an attended item

Florin Cutzu [a,*], John K. Tsotsos [b]

[a] Department of Computer Science, Indiana University, Bloomington, IN 47405, USA

[b] Department of Computer Science and Centre for Vision Research, York University, Toronto, Ont., Canada M3J 1P3

Received 23 October 2000; received in revised form 14 October 2002

Abstract

The selective tuning model [Artif. Intell. 78 (1995) 507] is a neurobiologically plausible neural network model of visual attention. One of its key predictions is that to simultaneously solve the problems of convergence of neural input and selection of attended items, the portions of the visual neural network that process an attended stimulus must be surrounded by inhibition. To test this hypothesis, we mapped the attentional field around an attended location in a matching task where the subject's attention was directed to a cued target while the distance of a probe item to the target was varied systematically. The main result was that accuracy increased with inter-target separation. The observed pattern of variation of accuracy with distance provided strong evidence in favor of the critical prediction of the model that attention is actively inhibited in the immediate vicinity of an attended location.

Keywords: Attentional field; Inhibition; Neural network; Spatial attention; Visual attention

1. Introduction

Functionally, visual attention is a mechanism that renders tractable the computationally complex search procedures required by practical vision problems. In Tsotsos (1990) it was shown how attention can be precisely defined and rigorously treated in terms of computational complexity theory. An important conclusion of that work was that the presence of an attentional selection mechanism is a matter of theoretical necessity whenever a vision system (whether biological or artificial) is confronted with a visual task of real-world complexity. A neural network model of visual attention, the *selective tuning* (ST) model, was developed on this theoretical foundation (Tsotsos, Culhane, Wai, Davis, & Nuflo, 1995). ST has a biologically inspired architecture and was intended to be a functional model of primate visual attention. ST belongs to the class of spatial selection models of attention; what distinguishes ST from other models in this class is its prediction of an annular region of attentional suppression surrounding the facilitated attended location. The purpose of the research reported in this paper was to test this critical prediction of the ST model. The model is briefly described here, and in full detail in Tsotsos et al. (1995).

1.1. Background

1.1.1. The selective tuning model

Complexity analysis provides the formal foundation for the conclusion that attention must tune the visual processing architecture to permit task-directed processing (Tsotsos, 1990). ST takes two forms: spatial selection, realized by inhibiting task-irrelevant neural connections, and feature selection, realized by inhibiting the neurons that represent task-irrelevant features.

The role of attention is to select a subset of the input image and a corresponding path through the processing hierarchy such as to minimize any interfering signals. This is one of the major components of the ST model that distinguishes it from others: it addresses the three-dimensional nature of the attentional influences in the visual neural network and not only the effects of attention on the image. The visual processing architecture posited by the model is a pyramidal network of units receiving both feed-forward and feedback connections,

* Corresponding author.
E-mail addresses: florin@indiana.edu (F. Cutzu), tsotsos@cs.yorku.ca (J.K. Tsotsos).

0042-6989/03/$ - see front matter
PII: S0042-6989(02)00491-1

similar to that proposed by Olshausen, Anderson, and Essen (1993).

The processing of the visual input in the network involves three main stages. During the first stage, a stimulus is applied to the input layer and activity propagates along feed-forward connections toward the output layer. The result is the activation of an inverted sub-pyramid of units and connections. The response of each unit depends on its particular selectivities, and perhaps also on a top-down bias for task-relevant qualities.

During the second stage, a hierarchy of winner-take-all (WTA) processes (Koch & Ullman, 1985) is applied to the network in a top-down, coarse-to-fine manner. The first WTA process operates in the top layer and covers the representation of the entire visual field: it computes the unit(s) with the largest response in the output layer, that is, the global winner. In turn, the global winner activates a WTA amongst its input units in the layer immediately below. This localizes the largest response within the receptive field of the global winner. All of the connections of the visual pyramid that do not contribute to the winner are pruned (i.e., attenuated). This strategy of finding the winner within each receptive fields and then pruning away irrelevant connections, is applied recursively through the pyramid, layer by layer. Thus, the global winner in the output layer is eventually traced back to its perceptual origin in the input layer. The connections that remain (i.e., are not pruned) may be considered the pass zone of the attentional beam, while the pruned connections an inhibitory zone around that beam, as shown in Fig. 1.

During the third stage, the selected stimuli in the input layer re-propagate through the network, being processed by the same neurons but this time without distracting stimuli in each receptive field, as if they had been presented on a blank background. Note that there is no change in identity of the winning neurons in the output layer; the winner initially selected remains the winner but its value is refined by this process.

In a converging network architecture as is present in the model and as is apparent in the visual cortex the convergence of feed-forward pathways leads to signal interference as signals merge layer by layer. The inhibitory surround imposed by our hierarchical WTA process removes this interference and causes an attended stimulus to be processed as if it appears on a blank background as far as the receptive fields relevant to its processing are concerned. Signal interference also motivated the shifter circuit proposal in Anderson and Van Essen (1987); however their method dealt only with the removal of interference and its implementation in Olshausen et al. (1993) did not predict a suppressive surround nor was it consistent with widespread attentional modulation observed throughout the visual cortex.

Just like a human observer, the model can be provided with a spatial cue to indicate the location of a relevant stimulus. While being presented with a cue, the model determines the location of the corresponding most active units in the output layer, and retains this information in the form of a bias in favor of these units. When the subsequent test stimulus appears, this bias remains in place and influences the WTA processes in the next layer down.

1.1.2. The distribution of attention in visual space

We use the term *attentional field* to describe the dependency of the intensity of attention on visual field location. The intensity of attention at a certain location

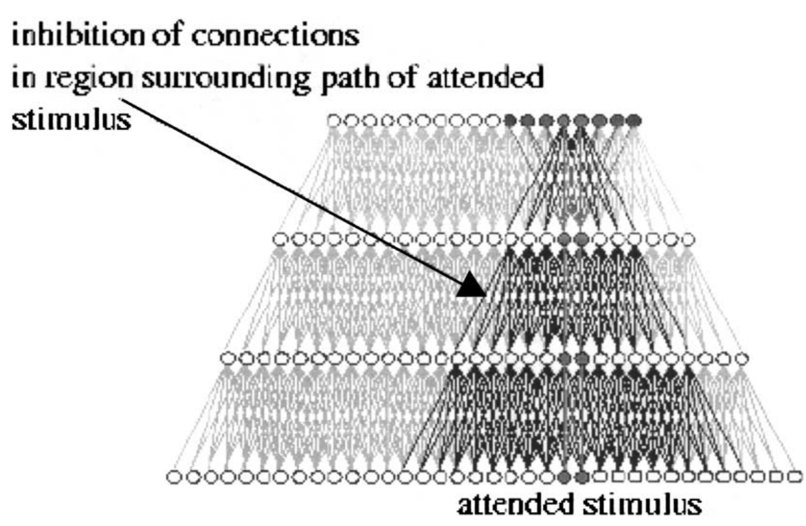

Fig. 1. A visual processing pyramid with four layers. Each unit is connected to seven units in the layer immediately above, as well as to seven units in the layer immediately below. The input layer (bottom layer) is numbered 1, while the output layer (top layer) is numbered 4. Note that feed-forward and feedback connections are not shown separately, instead, each reciprocal pair of connections is represented by a single line.

is defined as the capacity of the visual system (as measured by error rates or response times) to process a visual stimulus at that location.

According to the current models of spatial attention, namely *spotlight* (Eriksen & Hoffman, 1973; Posner, Snyder, & Davidson, 1980), *zoom lens* (Eriksen & St. James, 1986; Eriksen & Yeh, 1985) or *gradient* (Andersen & Kramer, 1993; Cheal, Lyon, & Gottlob, 1994), attention facilitates information processing in a connected region centered on the attended target, leaving the rest of the visual field essentially incapable to process visual input. The facilitated area is of fixed size and sharp boundary in the spotlight model, and has task-dependent size and spatial distribution in the zoom-lens model; in the gradient model the sharp boundary of the spotlight is replaced with a gradual decrease of the attentional field with distance from the attended target. Although these models disagree with regard to the size or shape of the area of perceptual facilitation they all assume that the attentional field decreases monotonically to zero with the distance from the focus of attention. In contrast, the attentional field of the ST network has an annular region of inhibition—i.e. the attentional field is negative—surrounding the region of perceptual facilitation centered on the attended target; outside the inhibited region, the model holds that the attentional field becomes positive again.

1.1.3. Psychophysical evidence for surround inhibition in spatial attention

Although the spotlight, zoom-lens and the gradient models and their variants dominate the current thinking about the spatial distribution of attention in the visual field, these schemes cannot explain a growing body of experimental evidence for attentional suppression effects in spatial vision. In this section we briefly describe several studies that have revealed suppressive effects in spatial attention.

In an early experiment, Skelton and Eriksen (1976) tested subjects with briefly presented displays consisting of eight letters equidistantly arranged on a ring centered on the fixation point. The subjects had to decide whether two precued letters (a 80 ms stimulus onset asynchrony interval was employed) were identical or different. Response time were fastest for adjacent and—surprisingly—diametrically opposite pairs of cued targets. The slowest response times occurred when one letter intervened between the two cued letters.

Kröse and Julesz (1989) measured the detectability of a randomly oriented T character appearing at a precued position in a ring of randomly oriented L characters at short presentation times (100 ms). In one experiment, the subjects had to simultaneously identify the letter at the precued position (a L or a T) and report an additional T character (the non-cued target) that could appear at a surprise location in the ring of Ls. The target was not more detectable when close to the cued location—in fact, the detectability of the non-cued target was found to be slightly higher at positions opposite the cued target for one subject and independent of the distance to the cued target for the other subject.

These two early studies thus revealed a degree of performance improvement with distance to the focus of attention, suggestive of attentional suppression. More recent experiments have provided stronger evidence for surround inhibition in selective attention.

Pan and Eriksen (1993) mapped the attentional field in two dimensions by using a same–different letter comparison task where an irrelevant, competing letter, acting as attentional probe, was presented in the display together with the target pair. By systematically varying target separation and the distance of the distractor letter to the target pair, the attended region was determined to be elliptical, with the major axis determined by the target pair, and dimensions determined by target separation. The authors interpreted the observed attentional spatial effects as the expression of the inhibitory field centered on the selected stimuli, rather than of a positive selection mechanism.

In a related study, using the same method of response competition by incompatible distractors, Eriksen and Pan (1993) probed the horizontal extent of the attentional field around a target. The extent of the task-relevant area was manipulated by varying the size of a figure in which the target was embedded. The distance between distractors and the borders of the figure was also manipulated. Interestingly, the interference from incompatible distractors varied inversely with the distance from the edges of the figure containing the target and did not depend on their distance from the target itself. Once again, the results were viewed as evidence for an inhibitory field surrounding the attended area.

Bahcall and Kowler (1999) measured the identification accuracy for two target letters placed in a ring of 24 letters. Attention was directed to the two target letters by precuing their locations. Identification error rate decreased with inter-target distance, indicating that attending to one spatial location results in a trade-off whereby the processing of neighboring stimuli is decreased. The distance effect was replicated with several types of attentional cues, and controls ruled out the involvement of sensory masking or sensory transients. The authors interpreted the results as expressing either reduced precision in attentional targeting or reduced processing capacity in the neighborhood of the attended location.

In the experiments of Caputo and Guerra (1998) the target, the distractor and the non-targets forms were arranged in a circular display. The target was the form singleton and the distractor was the color singleton. Subjects performed a length discrimination task on a line segment included in the target. In one experiment, the distance between target and distractor was manipulated. The authors found that discrimination performance improved with distractor–target distance, and interpreted this finding as evidence for surround inhibition in selective attention.

Cave and Zimmerman (1997) studied the allocation of attention in the visual field by using spatial probes in experiments where subjects searched for a target letter in eight-letter displays. Two of their conclusions are relevant to this discussion, namely that attentional strength is flexibly adjusted according to the confusability between target and distractors, and that distractor locations near the target receive more inhibition than those farther from the target.

In a related study, Kim and Cave (1999) had subjects search, in a circular array, for a square among circles shapes, ignoring color differences. Response times showed that the color singleton distractor could draw attention to its location only when located far from the target, thus suggesting local inhibitory interaction.

A series of experiments recently carried out by Mounts lends further support to the notion that an attended location is surrounded by an inhibitory region of limited spatial extent. In Mounts (2000a) an irrelevant but attentionally salient item was found to degrade subject performance in a form discrimination task, effect which decreased with the distance between the salient stimulus and the target. In Mounts (2000b), a task-irrelevant color or orientation singleton distractor was shown to affect form discrimination in a multi-element display. The effect decreased with the distance between the salient distractor and the target whose form was to be discriminated. Target detectability was on the other hand unaffected by the distractor, indicating that a genuine attentional—rather than sensorial effect—had been observed.

1.2. Objectives

The experimental studies reviewed in Section 1.1.3, though generally in agreement with the predictions of the ST model, do not conclusively confirm it, as they were not specifically designed to verify its predictions. For example, in the Bahcall–Kowler work both targets were cued; to properly test the ST model the attentional beam must be directed at a single stimulus (location) in the visual field. The experiments described in this paper were designed to provide a critical test for the ST model by contrasting its predictions to the predictions of the classical spatial attention models. Specifically, our goal was to determine whether the attentional field around an attended location exhibits the limited extent, excitatory center-inhibitory-surround pattern of variation predicted by the ST model, or rather the monotonic variation of the classical attention models.

2. Psychophysical experiments

2.1. Principle of the experimental method

The principle of our experimental method was to direct attention to a given reference location in the visual field, while concomitantly measuring attentional field intensity (visual information processing capacity) at various other, probe, locations. By systematically varying the reference–probe distance one can then measure the quantity of interest, the dependence of the attentional field intensity on distance to the focus of attention.

The experimental requirements were threefold. The first, obvious, requirement was that the visual task had to engage visual attention. The L–T letter discrimination task, known to require visual attention Julesz and Bergen (1983), was used. Discrimination accuracy was employed as a measure of the intensity of the attentional field.

Second, it was necessary to focus attention on a prespecified, known location in the visual field in order to establish a reference point around which to probe the attentional field. In terms of the ST model, this would cause the "pass zone" of the network to be anchored to the cued location. Precuing the reference location was used for this purpose. A peripheral, or direct, cuing method was employed in which the cue appeared briefly at the target location, shortly before the test image. Such cues draw attention automatically (Jonides, 1981) and act in a bottom-up fashion, presumably activating the so-called transient attentional subsystem (Egeth & Yantis, 1997; Nakayama & Mackeben, 1989; Weichselgartner & Sperling, 1987).

Third, it was necessary to ensure that differences in performance at different visual field locations were due to attentional effects and not to differences in retinal resolution. To obtain equal retinal resolution, the stimuli were arranged in a circular array display with fixation point in its center.

2.2. Outline of experiments

In all experiments, the subjects were required to detect target(s) among distractors following the presentation of a cue and the variable of interest was the dependence of performance—and thus of attentional field intensity—on target distance to the cued location.

Despite their common goal, the principle of the method of the two main experiments was quite different. Conceptually, the important experiments are Experiments 1 and 4; Experiments 2 and 3 were controls for Experiment 1. In Experiments 1, 2, and 3 the two targets differed in color from the distractors, the subject's task being to report whether they were identical or different. In Experiment 4 the sole target differed in shape from

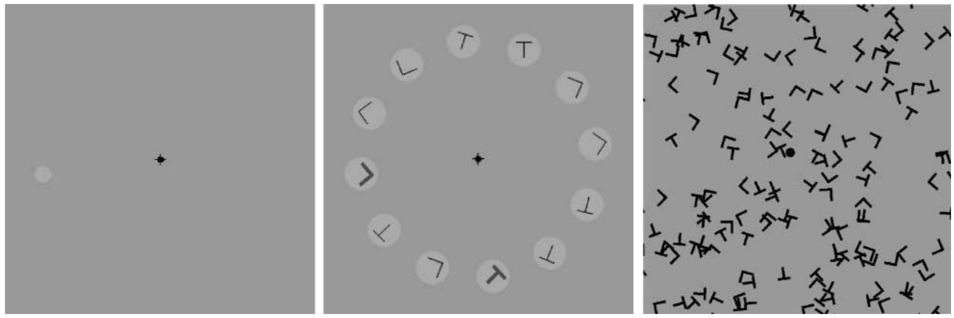

Fig. 2. Experiment 1. The basic trial sequence. (Left) The cue, a light gray disk indicated the position of the reference target character in the following, test, image. The cue was shown for 180 ms. (Middle) Test image, shown for 100 ms. The target characters were red (drawn in this figure with thick lines), the distractors were black (drawn with thin lines). The task was to decide whether the two targets are identical or different by pressing certain keys on the keyboard. (Right) The mask was shown until the subject responded.

the distractors, and the subject's task was simply to report its existence.

2.3. Experiment 1

2.3.1. Method

The experimental sequence, arranged in Fig. 2 from left to right, consisted of cue, test image, and mask.

The cue, a briefly displayed light gray disk, anticipated the position of the reference target in the following test image. The cue was shown for 150 ms, a time interval known to be within the range of effective peripheral cuing (Nakayama & Mackeben, 1989). The cue was valid in all trials.

The test image consisted of six randomly oriented L shapes and six randomly oriented T shapes, arranged in random order on a circle centered on a fixation point. The radius of the circle was 4° and character size was 0.6° visual angle. The circular arrangement ensured that all stimuli were perceived at equal retinal resolution. The characters were evenly spaced, and were overlayed on light gray disks as shown in Fig. 2, middle panel. These support disks were identical to the cue in terms of radius and color, thus rendering the transition cue–test image smooth, without perceptual transients that could disrupt the perception of the stimuli. Two of the characters, the reference target, which appeared at the cued location, and the probe target, which appeared at a surprise location, were red and the rest, the distractors, were black. The orientation of the line segment joining the reference target to the probe target character was randomly changed from trial to trial.

The subjects were instructed to respond to the test image as quickly and accurately as possible. A two-alternative forced-choice (2AFC) task was employed. The subjects decided whether the two target characters were the identical or different (ignoring orientation differences) by pressing one of two keys on the computer keyboard. The subjects were instructed to maintain fixation on the fixation point in the center of the ring.

After 100 ms the test image was replaced by a mask consisting of multiple randomly colored L and T characters, in random orientations, scattered on the screen. The role of the mask was to erase the iconic memory of the target letters in the test display. It was during the mask that subjects made their response. Once subjects responded the mask was removed and another experimental sequence was initiated.

Since there were 12 characters in the ring (six Ls and six Ts) the variable of interest, there were six values for inter-target separation, from one, when the two target characters were neighbors, to six, when diametrically opposite. Each of the six inter-target separations was tested eight times in the *identical targets* condition (LL or TT, four times each) and eight times in the *different targets* condition (LT). Thus, an experimental session thus consisted of 48 same target and 48 different target trials in random order.

Ten paid subjects, undergraduate and graduate students, naive as to the purpose of the experiment, were tested.

2.3.2. Results

Target discrimination accuracy was determined for each of the six target separation values for each subject data set. The data from the individual accuracy vs. inter-target separation plots are displayed jointly in Fig. 3 in the form of a boxplot (a more sophisticated variant of an errorbar plot). Accuracy improved with increasing inter-target separation, increasing from approximately 55% when the targets were immediate neighbors to about 83% when diametrically opposite.

An analysis of variance test was performed on the set of 10 individual accuracy rates. At each of the six inter-target separation values, 10 mutually independent observations (10 individual accuracy values) were available. ANOVA indicated that the six values of across-subjects mean accuracy were not all equal ($F = 35.85$, Prob $> F = 5.55e-16$).

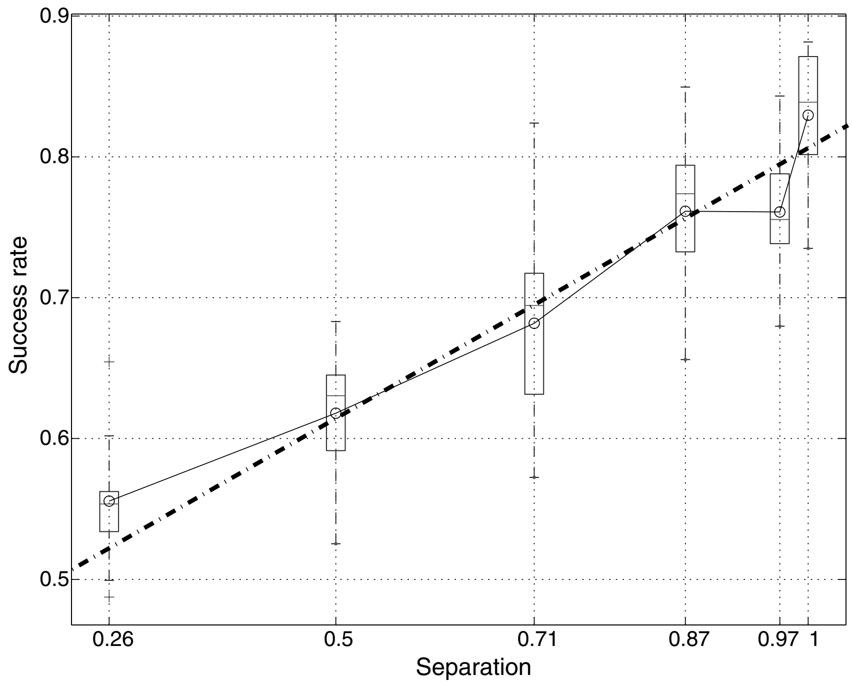

Fig. 3. Experiment 1. Boxplot of the dependence of target discrimination accuracy on target separation. The thick dash-dot line represents the linear regression model. Target separation is expressed in fractions of letter ring diameter. The boxes, one per separation value, have lines at the lower quartile, median, and upper quartile values. The whiskers (lines extending from each end of the box) show the extent of the rest of the data. Outliers, marked by the + symbol, are data with values beyond the ends of the whiskers. One of the two target locations was precued; the other target appeared at a surprise location on the ring of letters. Performance improved substantially with increasing inter-target distance.

A multiple means comparison test revealed which inter-target separation values resulted in different across-subjects mean accuracy rates. As detailed in Table 1, this test indicated that accuracy significantly improved with inter-target separation.

Thus, in the plot in Fig. 3, performance level at minimum inter-target separation (0.26) is significantly lower than at separations 0.71, 0.87, 0.97, and 1. Performance at inter-target separation 0.5 is significantly lower than at separations 0.87, 0.97 and 1. Performance is essentially the same at separations 0.87, 0.97 and 1, and is significantly greater than at separations 0.26, 0.5, 0.71.

Table 1
ANOVA results, Experiment 1

	1	2	3	4	5	6
1	0	0	1	1	1	1
2	0	0	0	1	1	1
3	1	0	0	1	1	1
4	1	1	1	0	0	0
5	1	1	1	0	0	0
6	1	1	1	0	0	0

The six rows and columns of the table correspond to the six inter-target separation values tested in the experiment. An entry of 0 indicates that the mean across-subjects accuracy rates at the two respective separations were not significantly different. An entry of 1 indicates a significant difference.

A linear regression analysis of the dependence of target discrimination accuracy on inter-target separation yielded the dash-dotted, positive-slope line plotted in Fig. 3, supporting the hypothesis that accuracy improved significantly, quasi-linearly with inter-target separation. The statistics of the linear regression were as follows. The R-square statistic $R^2 = 0.96$, indicating that the linear model accounted for 96% of the variability in the observations. The F statistic (for the hypothesis that all the regression coefficients are zero), and the p-value associated with F were $F = 91.5$ and $p = 0.0007$, indicating that it was extremely unlikely that all the regression coefficients were zero.

The global error rate for the pooled data was 29.9%, with a false different response rate (number of same pairs judged different, divided by number of trials) of 15.5% and a false same response rate (the number of different pairs judged same, divided by the number of trials) of 14.4.

The agreement among subject data was high, as indicated by the correlation coefficients between individual subject data (correct rates at different target separations) and the averaged subject data: 0.90, 0.92, 0.94, 0.93, 0.98, 0.96, 0.80, 0.96, 0.77, 0.90.

2.3.3. Discussion

The significant improvement of discrimination performance with inter-target separation observed in Experiment 1 suggested the existence of a zone of attentional suppression around the attended target. This finding is in concordance with the ST model, and in seeming disagreement with the predictions of the spotlight and related models. However, to eliminate potential confounding factors, two control experiments, Experiments 2 and 3, were run, as described below.

2.4. Experiment 2

In Experiment 1 target discrimination performance increased progressively with inter-target separation, the best performance occurring at the largest separation values. But what would be observed if inter-target separation would be increased even further? A simple extrapolation of the experimental performance-separation curves would suggest that increasing separation will improve performance. The purpose of Experiment 2 was to explore the effect of large inter-target separations on target discrimination. The setup of Experiment 1 was employed, the only modification being that the radius of the letter ring was increased by 150%, to 6°. Thus, the separation values were 1.5 times larger than in the previous experiment.

2.4.1. Method

The same methodology as in Experiment 1 was followed, only the diameter of the letter ring was increased. The sizes of the cue and letters were not changed. Five paid subjects, undergraduate students, were tested.

2.4.2. Results

Target discrimination accuracy was determined for each of the six target separation values for each subject data set. The individual accuracy vs. inter-target separation plots are displayed in Fig. 4 as a boxplot. Accuracy at first improved with increasing inter-target separation, increasing from approximately 58% correct rate, when the targets were immediate neighbors, to 75% correct rate, when two distractors intervened between the targets, then leveling off at this value for larger separations.

An analysis of variance test was performed on the set of five individual accuracy rates. At each of the six inter-target separation values, five mutually independent observations (five individual accuracy values) were available. ANOVA indicated that the six values of across-subjects mean accuracy were not all equal ($F = 15.77$, Prob $> F = 6.55e-07$).

A multiple means comparison test revealed which inter-target separations resulted in different across-subjects mean accuracy rates. As detailed in Table 2, the test revealed that initially accuracy improved with increasing separation, leveling off for three or more intervening letters.

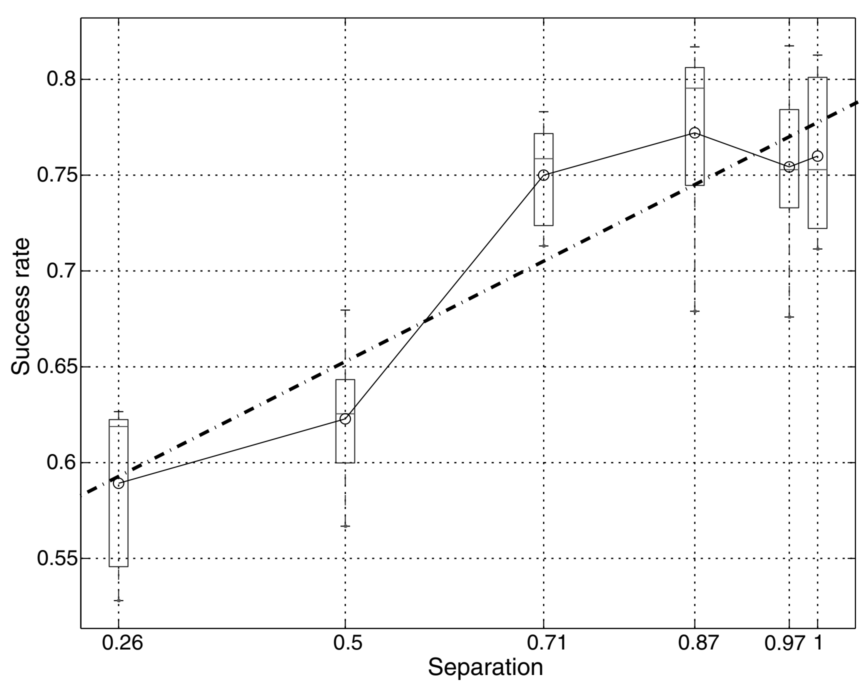

Fig. 4. Experiment 2. Boxplot of the dependence of target discrimination accuracy on target separation. The thick dash-dot line represents the linear regression model. Target separation is expressed in fractions of letter ring diameter. The boxes, one per separation value, have lines at the lower quartile, median, and upper quartile values. The whiskers show the extent of the rest of the data. Outliers (+) are data with values beyond the ends of the whiskers. One of the two target locations was precued; the other target appeared at a surprise location on the ring of letters. Performance initially improves with inter-target distance, reaches a maximum, and then levels off.

Table 2
ANOVA results, Experiment 2

	1	2	3	4	5	6
1	0	0	1	1	1	1
2	0	0	1	1	1	1
3	1	1	0	0	0	0
4	1	1	0	0	0	0
5	1	1	0	0	0	0
6	1	1	0	0	0	0

The six rows and columns of the table correspond to the six inter-target separation values tested in the experiment. An entry of 0 indicates that the mean across-subjects accuracy rates at the two respective separations were not significantly different. An entry of 1 indicates a significant difference.

Thus, in the plot in Fig. 4, performance level at minimum inter-target separation (0.26) does not differ significantly from performance at separation 0.5 and is significantly lower than at separations 0.71, 0.87, 0.97, and 1. Performance is essentially the same at separations 0.71, 0.87, 0.97 and 1, and is significantly greater than at separations 0.26 and 0.5.

A linear regression analysis of the dependence of target discrimination accuracy on inter-target separation yielded the dash-dotted, positive-slope line plotted in Fig. 4, supporting the hypothesis that accuracy improved significantly with inter-target separation. The statistics of the linear regression were as follows. $R^2 = 0.87$, indicating that the linear model accounted for 87% of the variability in the observations. The F statistic and the p-value associated with F were $F = 26.7$ and $p = 0.0067$, indicating that it was highly unlikely that all the regression coefficients were zero.

The global error rate for the pooled data was 30%, the false different response error rate was 15% and the false same response error rate was 15%.

The agreement among subject data was high, as indicated by the correlation coefficients between individual subject data (correct rates at different target separations) and the averaged subject data: 0.98, 0.85, 0.84, 0.95, 0.95.

2.4.3. Discussion

The leveling off of the discrimination performance (the plateau) observed at the three largest inter-target separation values indicates that the inhibitory ring has a finite extent. This seems to be in agreement with the ST model, according to which, once the probe target is located outside the inhibitory region, performance does not change significantly with distance to the attended location (the center of the inhibitory ring).

2.5. Experiment 3

ST predicts an inhibitory surround in the neighborhood of the focus of attention. To detect the inhibitory zone, one must know where the focus of attention is located. The role of the cue in the preceding experiments was to position the focus of attention at the location of the reference target. Given their special color, both targets pop out in the test display and compete for attention. Cuing was necessary to break this symmetry and direct attention to only one of the two targets. The purpose of Experiment 3 was to verify that the cue had indeed directed attention to the location of the reference target.

2.5.1. Method

The control consisted of cuing the center of character ring rather than the location of the reference target. The cue, being equidistant to all characters, provided no information about the location of the target elements functioning instead as a mere warning for the appearance of the test image. The test and mask images were identical to the ones in the peripheral cue condition. With the exception of the cuing method, the methodology (stimuli, procedure) followed was as described in Section 2.3.1.

Eight paid subjects, undergraduate and graduate students, naive as to the purpose of the experiment, were tested.

2.5.2. Results

Target discrimination accuracy was determined for each of the six target separation values for each subject data set. As shown in the errorbar plot Fig. 5, accuracy varied in the 0.5–0.6 interval, exhibiting no systematic dependence on separation.

An analysis of variance test was performed on the set of eight individual accuracy rates. Thus, at each of the six inter-target separation values, eight mutually independent observations (eight accuracy values) were available. ANOVA indicated that the six values of across-subjects mean accuracy were equal ($F = 0.36$, Prob $> F = 0.87$).

A linear regression analysis of the dependence of target discrimination accuracy on inter-target separation produced the dash-dotted, zero-slope line plotted in Fig. 5, indicating to that there was no linear dependence of accuracy on inter-target separation. The statistics of the linear regression were as follows: $R^2 = 0.05$, indicating that the linear model accounted for a negligible part of the variability in the observations; $F = 0.22$ and $p = 0.66$, indicating that the regression coefficients were zero.

The overall error rate was 41%, with a false different response rate of 22.25% and a false same response rate of 18.75%. Performance did not change significantly with distance between targets.

The agreement among subject data was quite low, as indicated by the correlation coefficients between individual subject data (correct rates at different target

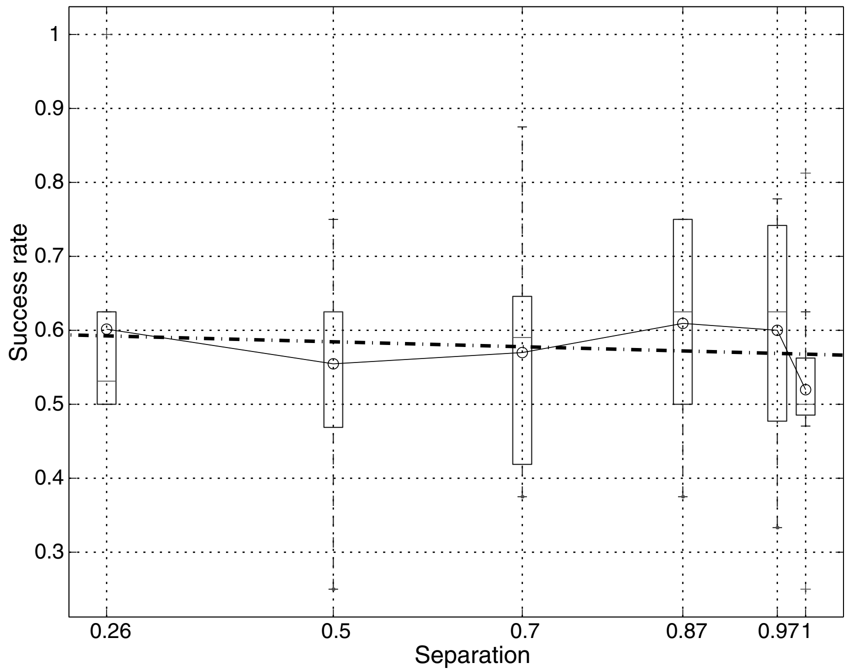

Fig. 5. Experiment 3. Boxplot of the dependence of target discrimination accuracy on target separation. The thick dash-dot line represents the linear regression model. Target separation is expressed in fractions of letter ring diameter. The boxes, one per separation value, have lines at the lower quartile, median, and upper quartile values. The whiskers show the extent of the rest of the data. Outliers (+) are data with values beyond the ends of the whiskers. In this control experiment the ring center (the fixation location) and not the target was precued. Performance did not change systematically with inter-target distance.

separations) and the averaged subject data: 0.13, 0.17, 0.65, 0.25, −0.39, 0.51, 0.47, −0.23.

2.5.3. Discussion

The systematic effect of inter-target separation on discrimination performance practically disappeared when the center of the character ring was cued rather than a location on its circumference. In addition, overall performance degraded in comparison to Experiments 1 and 2, where the cue had informative value relevant to the task; cuing the fixation cross has no informative value with regard to the location of the targets.

These results can be interpreted as follows. In Experiment 1, the cue centered attention on the reference target. As a consequence of the inhibitory-surround structure of the attentional field, discrimination performance increased with distance from the cued position resulting in improved performance for characters on the opposite side of the ring from the target. However, when the center of the display was cued all characters, being at equal distance from the focus of attention, were equally inhibited. This explains why variations in the distance between the two targets did not result in significant changes in discrimination accuracy.

2.6. Experiment 4

In Experiments 1–3 the targets were defined by their special, salient color. In this experiment, the design, while satisfying the requirements described in Section 2.1, was substantially simplified: the task involved only shape discrimination, and only one target element needed be detected. Specifically, the target was defined by its special shape, i.e., the target was the shape singleton in the display. The L–T discrimination task was again employed: the subjects were required to detect an odd L among T characters or an odd T among L characters arranged on a ring. The ring of characters was preceded by a briefly flashed cue, that did not always coincide with the position of the odd letter. We note that the goal of anchoring the attentional beam at the cued location can still be achieved even with an invalid cue, because briefly flashed, peripheral cues draw attention involuntarily, independently of their informational value Jonides (1981). The goal of this experiment was to study the detectability of the odd letter as a function of its distance to the cue.

The main advantage of the Experiment 4 over the previous designs is that, due to the fact that all letters have the same color and contrast, lateral masking among target and distractors is eliminated; the subject is confronted with a spatially uniform test display.

The simplification of the design came at a price, however: the task was much more difficult than in the previous experiments. Detecting an odd L among Ts (or vice-versa) is usually treated experimentally as a serial task, and subjects are given ample time to examine the display. However, in this experiment, brief presentation times (100 ms) were employed. To avoid a large number of false positives, the subjects were warned about the difficulty of task and were instructed to adopt a conservative strategy by responding "target" only when reasonably confident of detecting a different letter in the display.

2.6.1. Method

As in Experiment 1, a trial sequence consisted of cue, test image, and mask (see Fig. 6).

As in Experiment 1, the cue, a light gray disk presented for 180 ms, was identical to the disks over which the characters in the test display were overlayed. However, as opposed to Experiment 1, in this experiment the cue was only sometimes valid. The subjects were not informed about the presence of the cue.

The test image consisted of 12 black L and/or T characters in random orientations, overlayed on light gray disks, arranged on a circle centered on the fixation point. The dimensions of the ring and of the characters were the same as in Experiment 1. All characters had the same color, black. In 50% of trials there was no target: all the letters in the ring were identical (either all Ts or all Ls). In the other 50% trials the target was present: in 25% trials there was one L among 11 Ts (the case shown in Fig. 6) and in 25% trials there was one T among 11 Ls.

The effect of the cue on target detectability as a function of its distance to the target was the variable of interest in the experiment. In the target present trials, the target (the odd letter) was either at the position indicated by the cue (in which case the distance cue–target was zero), or at a different location on the ring. Since there were 12 letters in the ring, there were seven different values of the cue–target distance, corresponding to 0, 1, 2, 3, 4, 5, 6 intervening letters. Each target–cue distance value (0, 1, 2, 3, 4, 5, 6) was tested six times, resulting in $7 \times 6 = 42$ target present trials. There were 42 target absent trials. The orientation of the imaginary cue–target line was randomly changed from trial to trial.

The test image was shown for 100 ms and was followed by the mask image (identical to the one used in Experiment 1), which was shown for 2 s. During this time interval the subjects could respond to a detected target by pressing a key on the computer keyboard, or could wait for the next trial sequence if no target was detected. Following the removal of the mask, a new trial sequence was initiated. The subjects were instructed to maintain fixation at all times on the fixation point in the center of the ring.

Seven paid subjects, graduate and undergraduate subjects, were used in this experiment.

2.6.2. Results

The overall "target present" (hit) rate for the pooled data was 48%. The correct "no target" (reject) response rate was 69%. The false alarm rate, defined as the fraction of the non-target test images in which the subjects falsely detected a target was 31%. The false negative (miss) rate was 52%.

Task success rate (target detection accuracy, hit rate) was determined for each of the seven target–cue separation values for each subject data sets. The individual detection accuracy vs. inter-target separation plots are displayed in Fig. 7 in the form of a boxplot (obviously, the plot corresponds to the *target present* trials). As expected, detection accuracy was maximal at the cued location (zero separation), confirming that the effectiveness of the cue. Away from the cue, accuracy initially decreased but then increased again, peaking opposite to the cued location on the ring of letters.

An analysis of variance test was performed on the set of seven individual accuracy rates. Thus, at each of the seven cue–target separation values, seven mutually independent observations (seven individual accuracy values) were available. ANOVA indicated that the seven values of across-subjects mean accuracy were not all equal ($F = 13.45$, Prob $> F = 1.9e-08$).

A multiple means comparison test revealed which cue–target separations resulted in different across-sub-

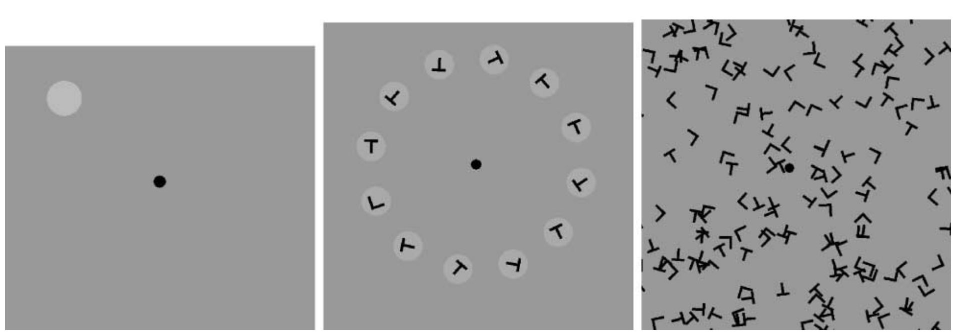

Fig. 6. Experiment 4. A typical trial sequence. (Left) The cue, a light gray disk, was shown for 180 ms. (Middle) Test screen, shown for 100 ms. The characters were overlayed on light gray disks identical to the cue. In this case the target is present: there is an odd L among the Ts in the test image. However, the cue is invalid. The subjects' task was to detect the odd letter in the ring. (Right) The mask was always removed after 2 s, whether the subject responded or not, and a new trial sequence was initiated.

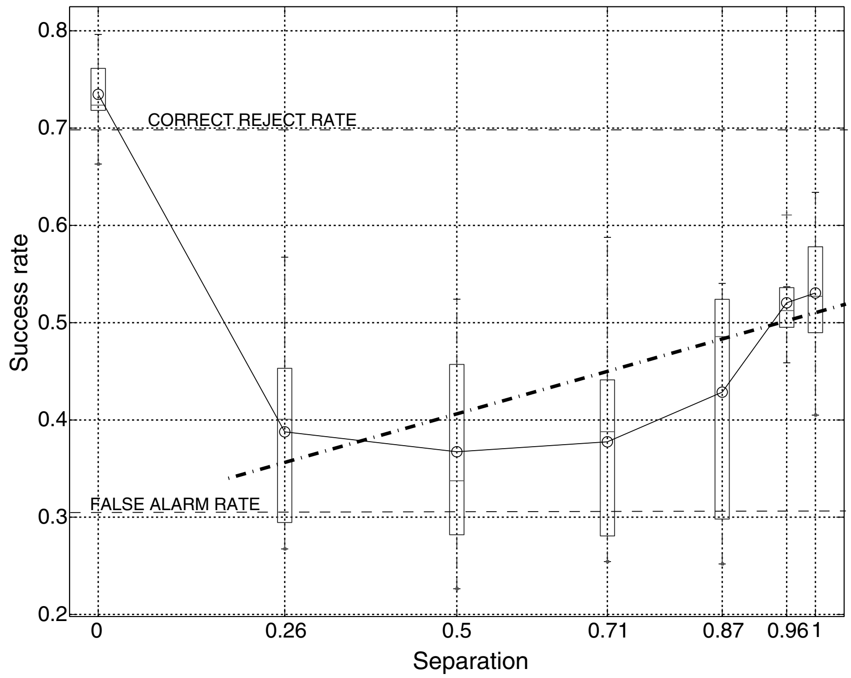

Fig. 7. Experiment 4. Boxplot of dependence of target detection accuracy on target–cue distance in the shape-singleton detection experiment. The cue–target distance was defined as the length of the chord joining the cued location on the ring to the target location on the ring. It was expressed in units of ring diameter, and ranges from 0 (the odd letter is at the cued location) to one diameter (the odd letter is diametrically opposite to the cued location). The interrupted lines represents the false alarm level (0.31) and the mean correct reject ('target absent') rates. The thick dash-dot line represents the linear regression model. The boxes, one per separation value, have lines at the lower quartile, median, and upper quartile values. The whiskers are lines extending from each end of the box and show the extent of the rest of the data. Outliers, marked by the + symbol, are data with values beyond the ends of the whiskers. Target detection performance peaks both at the cued location and diametrically opposite to it.

Table 3
ANOVA results, Experiment 4

	1	2	3	4	5	6	7
1	0	1	1	1	1	1	1
2	1	0	0	0	0	0	0
3	1	0	0	0	0	0	1
4	1	0	0	0	0	0	0
5	1	0	0	0	0	0	0
6	1	0	0	0	0	0	0
7	1	0	1	0	0	0	0

The seven rows and columns of the table correspond to the seven inter-target separation values tested in the experiment. An entry of 0 indicates that the mean across-subjects accuracy rates at the two respective separations were not significantly different. An entry of 1 indicates a significant difference.

jects mean accuracy rates. As detailed in Table 3, the test indicated that the accuracy at zero and maximal separation was significantly larger than at the other separations.

Thus, in the plot in Fig. 7, performance level at minimum cue–target separation (0) is significantly higher than at other separations. Performance at separation 0.5 is significantly below performance at minimum (0) and maximum (1) cue–target separation. Performance at maximum separation is inferior to performance at zero separation but superior to performance at separation 0.5.

A linear regression analysis of the dependence of target detection accuracy on target–cue distance for target locations not coinciding with the cue produced the dash-dotted, positive-slope line plotted in Fig. 7, supporting the hypothesis that detection improved in an approximately linear manner with cue–target distance. The statistics of the linear regression were $R^2 = 0.63$, $F = 6.7$ and $p = 0.05$.

It is noteworthy that target detection accuracy was greater that the false alarm error level for all cue–target distance values. This indicates that the subjects responded "target" only when very confident. Thus, most errors were of the false negative type: failure to detect the target.

Using the hit rates and false alarm rates, we computed the d' measure of detectability (Green & Swets, 1966) for each of the cue–target separation values. The false alarm rate was 0.31 (obviously, for target absent trials the cue–target distance was not defined). At the seven cue–target separations, the hit rates were 0.73, 0.39, 0.37, 0.38, 0.43, 0.52, 0.53, and the corresponding d' values, were, respectively, 1.12, 0.21, 0.16, 0.18, 0.31, 0.55, 0.57.

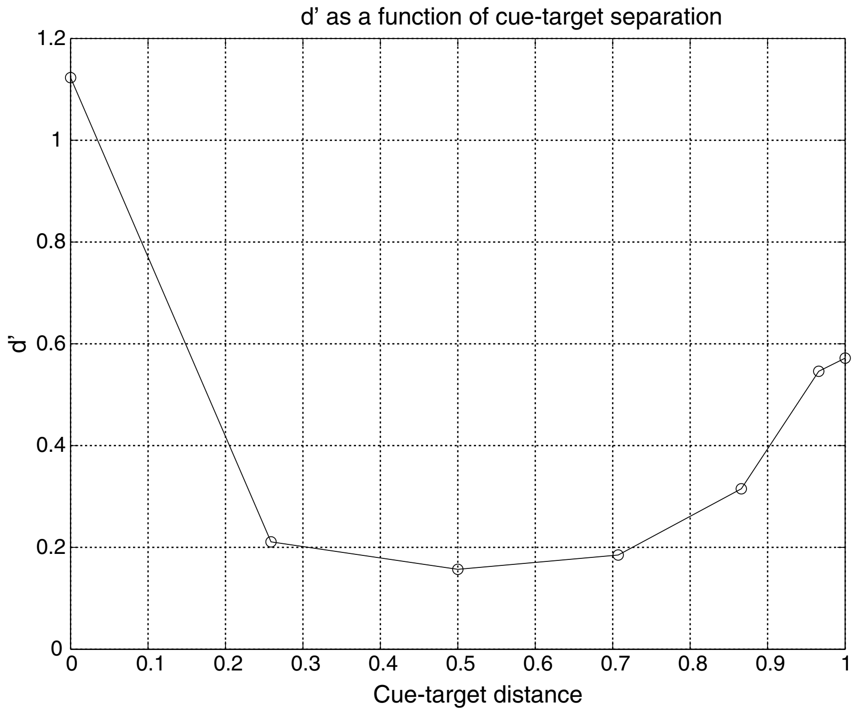

Fig. 8. Experiment 4. The d' measure of target detectability for each of the cue–target separations. d' peaks both at the cued location and diametrically opposite to it.

Fig. 8 displays the variation of the d' value with cue-singleton separation. Discriminability has two local maxima, one at the cued location, and the other at the furthest distance from the cued location.

The agreement among subject data was still high (but lower than in the other experiments), as indicated by the correlation coefficients between individual subject data (correct rates at different cue–target separations) and the mean subject data: 0.77, 0.81, 0.75, 0.94, 0.67, 0.98, 0.85.

2.6.3. Discussion

The U-shaped pattern of singleton detection variation observed in this experiment indicates the existence of an inhibitory zone, of annular shape and limited extent, surrounding the attended, facilitated target location.

Confirming the effectiveness of the cue, the detection of the singleton shape was most accurate at the cued location; in the previous experiments this facilitating cue effect could not be directly observed.

Experiment 4 is significant for an additional reason: the variation in the detectability of the odd letter can not be reasonably attributed to sensorial causes (lateral masking), as there were no differences in color or contrast among the characters in the display. These findings are similar to Bahcall and Kowler's, who also included conditions where there were no color differences, obtaining the same pattern of results. Thus, Experiment 4 makes a strong case for an attentional basis for the observed inhibitory surround.

In principle, the detection of the shape singleton in this experiment requires the identification of a minimum of two display items; thus, apparently, Experiment 4 is no different from the preceding experiments. In Experiment 4 however, the identity of the numerically dominant (distractor) letter was probably determined in a location-independent manner (say, by corroborating information available at a variety of locations in the visual field); the unique location of the singleton letter was the "probe" location where the intensity of the attentional field (centered on the cued location) was measured. In fact, in preliminary experiments the majority (distractor) letter was always "L"; therefore, the subjects were aware of the identity of the distractor letters before the test display was shown, and the task involve the detection of a single "T" target, not same–different discrimination. The results obtained in these preliminary experiments did not differ from the results obtained in the reported version of Experiment 4.

A potential problem with Experiment 4 is the relatively high false alarm rate. One may argue that subjects adopted a high criterion-strategy, as follows. Let us assume that, instead of having the U-shaped variation predicted by ST, target detectability simply declined with distance from the cue. Thus, if the subjects did not easily detect the target on a given trial, they would simply guess that the target was present, and located far away from the cue. This strategy would increase the number of target present responses for the furthest distances, even though the target might not actually have been detected. However, the behavior of the discriminability measure d' renders this scenario unlikely: target discriminability (however low) does increase for the furthest distances from the cue.

3. General discussion

The goal of the experiments reported in this work was the verification of the prediction of the ST model concerning the spatial distribution of attention in the vicinity of the focus of attention. We used a cuing procedure to anchor attention at a predetermined location in the visual field and a shape discrimination task to map the attentional field at various probe locations of equal retinal resolution.

We found that discrimination accuracy drops significantly in the immediate vicinity of the focus of attention only to increase again farther away from it. These results are rendered more plausible by the fact that the same pattern of variation was obtained with two rather different tasks: a same–different task run on two targets standing out from the non-targets by virtue their color, and a difficult shape-singleton detection task.

This finding directly contradicts the traditional view of the spatial distribution of attention which holds that performance degrades with distance to the focus of attention, either brusquely (as predicted by the spotlight model) or gradually (as predicted by the gradient model).

In our experiments at least one target located outside the focus of attention must have been identified for the task to be performed. We note, however, that this does not imply multiple foci of attention. According to the ST model, stimuli at non-attended (non-precued) locations are also processed, albeit to a reduced degree, proportional to the intensity of the attentional field there. In the framework of the ST model, the focus of attention (a salient or a precued location) is simply the location of the maximum of the attentional field—and thus it differs only quantitatively from the other locations in the visual field. The question can be raised whether the two red targets are processed serially or concurrently. The answer is that in the ST neural network the entire visual field is processed in parallel, even if not with the same "attentional intensity", given that only a subset of the neural network is facilitated.

Sensorial, non-attentional causes for the observed surround inhibition were rendered unlikely by control experiments. Crucially, accuracy ceased to vary systematically with inter-target separation when the center of the display—a location equidistant to the stimuli relevant to the task—was cued. Since the test images in the main and control experiments were identical, the difference in performance must be attributed to the differences in cuing, and likely represents, therefore, an attentional effect.

The objection can be raised [1] that by making the two targets different in color from the distractors the inhibitory effects can have a sensorial cause. It has been reported (Kooi, Toet, Tripathy, & Levi, 1994) that crowding (masking) effects are enhanced when the target and the flanker share more features. It is thus possible that the targets (both red) were more effective lateral masks for one another than were the distractors (which were black) for the targets.

In response, it must be pointed out that the inhibitory effects were not restricted to the situation where the two red targets were immediate neighbors; inhibition also occurred when there were intervening black letters between the two targets. In addition, in preliminary, unreported experiments where one of the targets was red and the other blue or green, and the distractors were black, the same inhibitory effect was observed.

Overall, our results confirm previous psychophysical studies, such as those reviewed in Section 1.1.3 that have demonstrating inhibitory effects consistent with attentional suppression. However, what distinguishes our approach from these previous studies, and renders it significant, is that our experiments were specifically designed to verify the predictions of a theoretical model of attention (the SM model).

3.1. Evidence for attentional suppression from neuroscience

A potential problem with psychophysical evidence in general is that it is open to different interpretations. Fortunately for the case for attentional suppression, a series of recent neurobiological studies have produced *direct* evidence for inhibitory effects in spatial attention.

Schall and Hanes (1993) recorded from neurons in an area responsible for purposive eye movements (the frontal eye field) in rhesus monkeys engaged in a visual search task. It was found that these neurons initially respond equally to both targets and distractors located in their receptive fields. However, while the neuronal response to the target continued until the saccade to the target, the response to the distractors was suppressed, and more so when the target was closer to the receptive field of the neuron.

Somers, Dale, Seiffert, and Tootell (1999) employed functional MRI to study human subjects performing a visual discrimination task necessitating attention. The design of the task was based on a recent study Joseph, Chun, and Nakayama (1996) demonstrating that a visual task traditionally considered preattentive (spotting an orientation singleton) can be impaired by a concurrent rapid serial visual presentation task. Strong

[1] We thank an anonymous reviewer for pointing out this potential problem.

attentional modulation was observed both in V1 and in the extrastriate areas. The activation at retinotopic stimulus representations was significantly increased when attention was directed to the location of that stimulus, being replaced with suppression when attention was directed elsewhere. Although the observed attentional modulation was determined to be spatially specific, the authors ruled out a spotlight interpretation, favoring an object selection mechanism over spatial selection.

Vanduffel, Tootell, and Orban (2000) used a double-label deoxyglucose technique to compare the metabolic activity changes occurring in the early stages of the macaque visual system induced during two visual tasks, identical perceptually and behaviorally, but different with respect to attentional demands. Attention-dependent changes were observed both in the LGN and V1. The observed attentional modulation was expressed as a retinotopically specific band of suppressed deoxyglucose uptake located peripheral to the representation of the attended stimulus. This represents direct evidence for attention-dependent suppression of irrelevant stimulus representations in an annulus surrounding the attended item.

3.2. Open questions

Aside from the details that remain to be filled in to derive a full picture of attentional suppression in retinal space (for example, the precise shape and size of the inhibitory region, does it extend in depth, how it varies with cue and target size, etc.), a number of more fundamental issues arise. One such question is whether attentional suppression (and facilitation) takes place in retinal space (as our studies and others indicate) or whether it exists in a reference system fixed with respect to the external world and not the retina. A even more fundamental question is whether attentional suppression takes place also in different feature spaces, such as spatial frequency, orientation, color, shape space, etc. We plan to address all these questions in our future research.

4. Conclusions

The ST model was derived in a first principles manner. The major contributor to those principles derives from a series of formal analyses performed within the theory of computational complexity, the most appropriate theoretical foundation to address the question "why is attention necessary for perception?" The model not only displays performance compatible with experimental observations but also does so in a self-contained manner. That is, input to the model is a set of real, digitized images and not preprocessed data. The predictive power of the model seems broad. An early prediction (Tsotsos, 1990) was that attention seems necessary at any level of processing where a many-to-one mapping of neurons was found. Further, attention occurs in all the areas in concert. The prediction was made at a time when good evidence for attentional modulation was known for area V4 only (Moran & Desimone, 1985). Since then, attentional modulation has been found in many other areas both earlier and later in the visual processing stream, and that it occurs in these areas simultaneously (Kastner, De Weerd, Desimone, & Ungerleider, 1998). Vanduffel et al. (2000) have shown that attentional modulation appears at early as the LGN. The prediction that attention modulates all cortical and perhaps even subcortical levels of processing has been borne out by recent work from several groups (e.g., Brefczynski & DeYoe, 1999; Gandhi, Heeger, & Bnoyton, 1999; Somers et al., 1999). The notions of competition between stimuli and of attentional modulation of this competition were also early components of the model (Tsotsos et al., 1995) and these too have gained substantial support over the years (Desimone & Duncan, 1995; Kastner et al., 1998; Reynolds, Chelazzi, & Desimone, 1999).

Crucially, the ST model predicts an inhibitory surround that impairs perception around the focus of attention (Tsotsos et al., 1995). This too has recently gained support, both psychophysical and neurobiological, (Bahcall & Kowler, 1999; Caputo & Guerra, 1998; Vanduffel et al., 2000). The existence of the inhibitory surround has been confirmed by the experiments presented in this paper, which were specifically designed to verify the predictions of the ST model.

The model further implies that preattentive and attentive visual processing occur in the same neural substrate, which contrasts with the traditional view that these are wholly independent mechanisms. This point of view has also been gaining ground recently (Joseph et al., 1996; Yeshurun & Carrasco, 1998). A final prediction is that attentional guidance and control are integrated into the visual processing hierarchy, rather than being centralized in some external brain structure. This implies that the latency of attentional modulations decreases from lower to higher visual areas, and constitutes one of the strongest predictions of the model. Additional predictions of the ST model are the spatial and temporal modulations of visual cortical responses around the focus of attention, and the existence of a WTA circuit connecting cortical columns of similar tuning. The ST model offers a principled solution to the fundamental problems of visual complexity, a detailed perceptual account of both the guidance and the consequences of visual attention, and a neurally plausible implementation as an integral part of the visual cortical hierarchy. Thus, the model "works" at three distinct levels—computational, perceptual, and neural—and offers a more concrete account, and far more specific predictions, than previous models limited to one of these levels. We are working to extend the model in several directions, and are particularly interested in seeing how its architecture might map onto the actual neural circuitry of visual cortex.

Acknowledgements

This paper has greatly benefited from the suggestions of Dr. Eileen Kowler and of two anonymous reviewers. We are grateful for their help.

References

Andersen, G. J., & Kramer, A. F. (1993). Limits of focused attention in three-dimensional space. *Perception and Psychophysics, 53*, 658–667.

Anderson, C. H., & Van Essen, D. C. (1987). Shifter circuits: a computational strategy for dynamic aspects of visual processing. *Proceedings of the National Academy of Science, 84*, 6297–6301.

Bahcall, D. O., & Kowler, E. (1999). Attentional interference at small spatial separations. *Vision Research, 39*, 71–86.

Brefczynski, J., & DeYoe, E. A. (1999). A physiological correlate of the 'spotlight' of visual attention. *Nature Neuroscience, 2*, 370–374.

Caputo, G., & Guerra, S. (1998). Attentional selection by distractor suppression. *Vision Research, 38*(5), 669–689.

Cave, K. R., & Zimmerman, J. M. (1997). Flexibility in spatial attention before and after practice. *Psychological Science, 8*, 399–403.

Cheal, M., Lyon, D. R., & Gottlob, L. R. (1994). A framework for understanding the allocation of attention in location-precued discrimination. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 47*, 699–739.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48*, 269–297.

Eriksen, C., & Pan, K. J. B. (1993). Attentional distribution in visual space. *Psychological Research, 56*(1), 5–13.

Eriksen, C. W., & Hoffman, J. E. (1973). The extent of processing of noise elements during selective encoding from visual displays. *Perception and Psychophysics, 14*, 155–160.

Eriksen, C. W., & St. James, J. D. (1986). St. Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics, 40*, 225–240.

Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 583–597.

Gandhi, S. P., Heeger, D. H., & Bnoyton, G. M. (1999). Spatial attention affects brain activity in human primary visual cortex. *Proceedings of the National Academy of Science, 96*, 3314–3319.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye movement. In J. B. Long & A. D. Baddeley (Eds.), *Attention and performance IX* (*Vol. IX*, pp. 187–203). NJ, Erlbaum: Hillsdale.

Joseph, J. J., Chun, M. M., & Nakayama, K. (1996). Attentional requirements in a "preattentive" feature search task. *Nature, 379*, 805–807.

Julesz, B., & Bergen, J. R. (1983). Textons, the fundamental elements in preattentive vision and perception of textures. *Bell System Technical Journal, 62*, 1619–1645.

Kastner, S., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1998). Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science, 282*, 108–111.

Kim, M., & Cave, K. R. (1999). Top-down and bottom-up attentional control: On the nature of interference from a salient distractor. *Perception and Psychophysics, 61*, 1009–1023.

Koch, C., & Ullman, S. (1985). Selecting one among the many: a simple network implementing shifts in selective visual attention. *Human Neurobiology, 4*, 219–227.

Kröse, B. J. A., & Julesz, B. (1989). The control and speed of shifts of attention. *Vision Research, 29*(11), 1607–1619.

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral-vision. *Spatial Vision, 8*(2), 255–279.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229*, 782–784.

Mounts, J. R. W. (2000a). Attentional capture by abrupt onsets and feature singletons produces inhibitory surrounds. *Perception and Psychophysics, 62*, 1485–1493.

Mounts, J. R. W. (2000b). Evidence for suppressive mechanisms in attentional selection: Feature singletons produce inhibitory surrounds. *Perception and Psychophysics, 62*, 969–983.

Nakayama, K., & Mackeben, M. (1989). Sustained and transient components of focal visual attention. *Vision Research, 29*, 1631–1647.

Olshausen, B. A., Anderson, C. H., & Essen, D. C. V. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *Journal of Neuroscience, 13*, 4700–4719.

Pan, K., & Eriksen, C. W. (1993). Attentional distribution in the visual field during same–different judgments as assessed by response competition. *Perception and Psychophysics, 53*, 134–144.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*, 160–174.

Reynolds, J., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas v2 and v4. *Journal of Neuroscience, 19*, 1736–1753.

Schall, H. B., & Hanes, D. P. (1993). Neural basis of saccade target selection in frontal eye field during visual search. *Nature, 366*, 467–469.

Skelton, J. M., & Eriksen, C. W. (1976). Spatial characteristics of selective attention in letter matching. *Bulletin of the Psychonomic Society, 7*(2), 136–138.

Somers, D. C., Dale, A. M., Seiffert, A. E., & Tootell, R. B. (1999). Functional mri reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Science, 96*, 1663–1668.

Tsotsos, J. K. (1990). Analyzing vision at the complexity level. *Behavioral and Brain Sciences, 13*, 423–469.

Tsotsos, J. K., Culhane, S. W., Wai, Y. L., Davis, N., & Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence, 78*, 507–547.

Vanduffel, W., Tootell, R. B. H., & Orban, G. A. (2000). Action-dependent suppression of metabolic activity in the early stages of the macaque visual cortex. *Cerebral Cortex, 10*, 109–126.

Weichselgartner, E., & Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science, 238*, 778–780.

Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature, 396*, 72–75.